# Chapter 9
# Reflection and Impulse as Determinants of Human Behavior

**Anand Krishna and Fritz Strack**

## The Missing Link Between Knowledge and Action

The experience of being and acting human is a universal theme of cultural expression. From movies to great literary works, great weight is placed on both of these aspects, being and acting. Literary figures that resonate in our minds do so both because of their heroic (or antiheroic) actions and their inner life—their motivations, feelings, and thoughts on their own existence and actions. Sometimes this inner life reflects a clever, reflective type of thinking, as with the Danaans in Homer's *Iliad* and their plan to defeat the Trojans by infiltrating the city in a wooden horse. But equally fascinating are the stories of protagonists who show an impetuous, impulsive side, such as Icarus, the Greek youth who flew too close to the sun and thereby melted his artificial wings despite his father's specific warning. Lastly, there are those figures who are faced with an inner conflict between reflection and impulse—Odysseus, for all his cleverness, cannot resist the temptation to hear the deadly song of the Sirens, yet his forethought in having himself tied to the mast of his ship defeats the impulse to drown himself trying to reach them.

These three Greek legends appear to show different ways of thinking that lead to different results. Coming up with complex plans such as the Trojan Horse requires reflection and active use of knowledge, whereas such deliberation would seem antithetical to Icarus' flight. However, these two different modes of thought do seem to come together in a fashion, as is in Odysseus' use of forward planning to defeat the consequences of his impulsive decision. In everyday life, as well, people often approach situations in these two ways—spending a while considering what to eat for lunch and making a deliberate selection but also automatically grabbing a sweet dessert without thinking about it at all. When considering how these mental

A. Krishna (✉) • F. Strack
Institute for Psychology, University of Würzburg, Würzburg, Germany
e-mail: krishna@psychologie.uni-wuerzburg.de



P. Meusburger et al. (eds.), *Knowledge and Action*, Knowledge and Space 9,
DOI 10.1007/978-3-319-44588-5_9

processes work, it makes sense to look at the two styles of thinking separately to gain an understanding of their interaction and of their actual effect on behavior.

In this chapter we seek to illuminate the characteristics of these two processes, show their interactions with each another, and point out their common effect on behavior. To do so, it is first necessary to evaluate each system independently, examining historical and current perspectives on reflective and impulsive styles of thought. Thereafter, we present an integrative model of thinking and action in an endeavor to identify when which system of thought will be active and under what circumstances it will influence behavior.

## A Short Overview of Reflective and Impulsive Styles of Thinking

### *Theories of Reflection*

The idea that human behavior is based on active, reflective thought guided by the principle of attaining beneficial things is old and makes intuitive sense. It is difficult to argue why people would actively decide to act in a fashion that they know is bad for them without some belief that the action would ultimately be positive. In this conception of human thought, negative outcomes can be explained by a lack of information. The Greek philosopher Socrates, for example, proposed that people would otherwise act in ways that were good for them.

From a social psychological perspective, this kind of thinking is exemplified in expectancy-value theories and the concept of *homo oeconomicus* (e.g., Fishbein & Ajzen, 1975). The theory of planned behavior (Ajzen, 1985) is an established example of an expectancy-value model (Conner & Armitage, 1998). It depicts behavior as a function of several specific mental factors. In this conceptualization the three determinants of behavior are the attitude toward the behavior, the subjective norm relevant to the behavior, and the perceived behavioral control over the behavior. An attitude toward a specific behavior is generated by multiplying the evaluation of a possible perceived outcome of the behavior (a value) by the perceived likelihood of that outcome (an expectancy) and then summing the results of this multiplication for all possible outcomes. Similarly, the subjective norm is calculated by multiplying the actor's motivation to comply with another person's expectation by the perceived likelihood that that person holds that expectation over all persons. By contrast, perceived behavioral control is a function of the perceived power of behavior-inhibiting or behavior-facilitating factors multiplied by the likelihood that the actor has access to these factors.

The assumption in the theory of planned behavior is, therefore, that a human actor's calculation of these three determinants of behavior is optimally based on all available information. Once the determinants are established, the actor will integrate

them for all possible behaviors and select the best option. This behavior is then initiated via a behavioral intention.

This idea is echoed in various domains, both historically and in more modern contexts, such as organizational psychology (Vroom, 1964), addiction research (Sutton, Marsh, & Matheson, 1987), and education (Wigfield & Eccles, 2000). There is solid empirical support for the ability of expectancy-value models to predict intentions and behavior (e.g., Armitage & Conner, 2001). However, even the strongest empirical studies do not conclude that this kind of thinking can completely predict behavior. In particular, it seems unlikely that behavior occurring without conscious thought could be dependent on this kind of deliberative, intentional processing (e.g., Langer, Blank, & Chanowitz, 1978). Therefore, one must consider alternative mental processes that are characterized by less deliberative processing.

## *Theories of Impulse*

Expectancy-value models can describe the careful planning behind the Trojan horse well, but they seem less able to explain why Icarus would immolate his wings. The kind of action implied in the latter myth is apparently not influenced by the knowledge of the action's expected consequences. It may rather be seen as reckless, as based on an impulse that seems to instigate behavior automatically. This determinant of behavior differs from the rational assumptions of Socrates or the theory of planned behavior (Fishbein & Ajzen, 1975) and must be explained by other means. Historically, this point is recognized by Aristotle in the *Nicomanchean Ethics* (Crisp, 2000) when he argues that people may at times act against their judgment. In his view an overwhelming passion—physical feelings (e.g., hunger) and emotions (e.g., fear or pleasure)—directly implies a practical conclusion to act on it. This practical conclusion may at times overpower the conclusions reached by reason.

Psychology as a discipline has focused mostly on factors that affect the ability of reasoned conclusions to control passionate ones (Hofmann, Friese, & Strack, 2009). For example, children's ability to forgo a small, immediate reward in favor of a larger, delayed reward has been shown to depend on various factors, including opportunity to pay attention to the immediate reward and thinking happy or sad thoughts (Mischel, Ebbesen, & Raskoff, 1972). Muraven and Baumeister (2000) conceptualize the control of impulses as being achieved by a metaphorical "muscle" that is powered by limited resources and is subject to fatigue, a state called *ego depletion*. In this view, reasoned control of impulses will fail if remaining resources of self-control are insufficient to overcome impulse strength. Personality variables have been linked to the ability to control impulses (e.g., Block & Block, 1980; Carver, 2005), as have physiological variables such as blood glucose and alcohol levels (Bushman & Cooper, 1990; Gailliot et al., 2007) and situational factors such as the availability of tempting stimuli (Schachter, 1971).

Another approach to impulses and their effect on behavior is provided by research on implicit attitudes. Implicit attitudes are seen as spontaneous, automatic affective and behavioral responses to attitude objects (Greenwald & Banaji, 1995). The view that impulsive behavior comes from overwhelming passion caused by a situation maps well onto this idea that attitude objects may spontaneously cause affect and behavior via automatic processes. This view contrasts with the position that attitudes are evaluations of target objects (e.g., Eagly & Chaiken, 1993). The evaluative position implies a more deliberative assessment of object properties, which corresponds to processes similar to those described in the theory of reasoned action (Fishbein & Ajzen, 1975). Researchers studying implicit attitudes generally adopt indirect attitude measures such as the affective priming paradigm (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) or the implicit association test (IAT) (Greenwald, McGhee, & Schwartz, 1998). These instruments typically use a combination of valent and categorical stimuli and measure their inhibition or facilitation of a target behavior. Affective priming studies show that people categorize evaluative adjectives such as *good* or *bad* faster when they are congruent in valence to an attitude object shown immediately beforehand (Fazio, 2001). The IAT measures the difference between reaction times when participants are asked to categorize an object by attitudinal categories via a key-press reaction also associated with a particular valence (Greenwald et al., 1998). These measures show that behavioral responses to specific stimuli are influenced by the valence of these stimuli. The difference between what these implicit measures capture and what traditional evaluative measures (e.g., self-report questions) assess is apparent from their only moderate correlations (Fazio & Olson, 2003) and by the relative robustness of implicit measures with regard to participants' conscious control (Banse, Seise, & Zerbes, 2001; Kim, 2003; but see Mierke & Klauer, 2001, and Fiedler & Bluemke, 2005, for critical discussions). One explains these results by stating that the spontaneous response to an attitudinal object is governed by associations that may differ in content from the results of a deliberative evaluation. These associations occur between perceived objects, behavior, and affect. The link between impulse and behavior has some theoretical basis—including the ideomotor principle (James, 1890) and the perception-behavior link (Bargh, 1997)—as does a direct link between perception and affective response (Zajonc, 1980). However, research on implicit attitudes has not succeeded in explaining what an implicit attitude actually is, beyond the tautological functional definition that an implicit attitude is what is measured by indirect measures (Strack & Deutsch, 2007). With evidence accruing that implicit attitudes may be strongly affected by the context (Wittenbrink, Judd, & Park, 2001) and type (Bosson, Swann, & Pennebaker, 2000) of the measurement, the question of what exactly an impulse might be is becoming ever more relevant.

## Bridging the Gap: The Reflective-Impulsive Model

The research and models presented so far shed some light on the mysteries of human behavior. The careful plan behind the Trojan Horse seems rooted in deliberative, expectancy-value thinking, whereas Icarus' doomed flight into the sun might be seen as a lack of self-regulatory resources preventing him from automatically flying close to the bright sun. What has yet to be addressed is the question of the interaction of these systems—how does Odysseus defeat the Siren's call, or in more general terms, how do deliberative, rational thought and impulsive affect and action interact? In the past three decades many researchers have attempted to address this question. In the realm of social psychology, the challenge has been met with many different dual-process models (see Chaiken & Trope, 1999, for examples). These dual-process models stem from research on different topics, including persuasion (Chaiken, 1987; Petty & Cacioppo, 1986), stereotyping (Devine, 1989), causal attribution (Gilbert, Pelham, & Krull, 1988), and general cognition (Epstein, Lipson, Holstein, & Huh, 1992). Although these models deal with different aspects of human behavior and cognition, they share certain structural traits. In all cases, two modes of information processing are proposed, one of which is characterized by effortful, rule-based processing, the other by almost effortless, associative processing. This distinction is delineated clearly by Sloman (1996), who describes both of these modes in detail.

Many influential dual-process models conform to this structure, but not all models provide for interactions between the proposed processes (for an overview, see Smith & DeCoster, 2000). Even fewer models make statements about how both processes influence behavior and whether they are interdependent in doing so. Rather, they specify how information is processed at a cognitive level—they describe how people think about what to eat for lunch and how they process the information that sweet desserts are available, but these models do not go beyond the formation of a behavioral decision on what would be best to eat. They also cannot describe how behavior can be instigated independently from a decision (e.g., Strack & Neumann, 1996)—how it can be that one decides to eat a healthy lunch but still reach out for the sweet muffin. In order to address these issues and fill the gap between knowledge and action, an integrative model that incorporates both elements is needed. For this purpose, Strack and Deutsch (2004) propose the reflective-impulsive model (RIM), a dual-systems model conceived to clearly define and integrate the theoretical perspectives so far presented under the headings of impulsive and reflective.

### *The Reflective System*

The RIM refers to the rational, rule-based system of thought as the reflective system. Slow and driven by resources of working memory, this system has limits on its capacity for information processing. It may be disengaged from processing under certain circumstances, but it is capable of generating knowledge via syllogistic inferences. It does so by activating concepts and possible relations between these concepts and then assigning a truth value to the proposition formed by the concepts and the relation. For example, the concepts *salad* and *health* may be activated, along with the relation of *is/is not*. In this case the reflective system would assign the value of *true* to the configuration, yielding the proposition *salad is healthy*. If the concept of *muffin* was activated instead of *salad*, the truth value might instead be *false*, yielding the proposition *muffin is not healthy*.

It is important to note that any other relation might be used beyond *is/is not*, such as *implies/does not imply*, *causes/does not cause*, or *is/is not a member of*, to name a few. If several related propositions are constructed, new knowledge may be generated by the combination of these propositions. In keeping with the example propositions given above—*salad is healthy* and *muffin is not healthy*—an additional proposition of *being healthy is good* might yield both the conclusion that salad is good and that muffins are not. Because the reflective system is able to assign truth values to statements, rule-based inferences can be drawn in order to maximize the consistency of the resulting representation (Gawronski & Strack, 2004). The ability of this mode of processing to help a person generate and infer conclusions makes it extremely flexible and useful for facilitating many operations typically associated with deliberative thought, including expectancy-value judgments and advanced social behavior (e.g., the discernment of people's states of belief; see Wimmer & Perner, 1983, for example) and the learning of new connections between concepts without much or any repetition.

The reflective system is limited by boundary conditions that constrain its ability to process information. Two of the most important of these conditions are working memory capacity and arousal. The activation of concepts and relations and the transformation of the resulting propositions are assumed to take place in the working memory (Baddeley, 1986). These dynamics provide a functional limit to the complexity and scope of reflective operations, in that the capacity of working memory may be insufficient to contain all the required propositions for a given operation concurrently. This statement is bolstered by studies showing that an impairment of working-memory capacity through a manipulation of cognitive load impairs logical reasoning (De Neys, 2006; DeWall, Baumeister, & Masicampo, 2008). Arousal, by contrast, affects reflective processing in a nonlinear fashion resembling the Yerkes-Dodson Law (Yerkes & Dodson, 1908)—intermediate levels of arousal facilitate the operation of the reflective system. Evidence exists that high levels of arousal reduce complexity in social judgments (Baron, 2000; Lambert et al., 2003; Paulhus & Lim, 1994), whereas low arousal, characteristic in a state of fatigue, for example, is also associated with lowered capacity to engage in reflective processing.

Important and ubiquitous cognitive phenomena rely upon the reflective system's ability to assign truth values to relations between concepts, an example of which is negation. The RIM predicts that negations of propositions can be processed only under circumstances in which the reflective system can be engaged, that is, under conditions endowed with resources sufficient for efforts to engage in processing. This statement differs from models based on the assumption that negations may be "tagged" onto propositions once and for all and henceforth no longer require reflective engagement (e.g., Gilbert, 1991), and it is supported by evidence that negations require cognitive resources to process (Wason, 1959).

Reflective processing is accompanied by a state of noetic awareness of whether something is the case or not. This awareness may sometimes be accompanied by a particular feeling that is processed consciously, a state of experiential awareness. It should be noted, however, that some operations in the reflective system may require so few resources that they can be processed without corresponding noetic awareness or a feeling of intentionality (Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009).

## *The Impulsive System*

Aside from the reflective system, the RIM proposes the existence of an associative system of thought called the impulsive system. This system works continuously and effortlessly. Whereas the reflective system uses working memory capacity, the impulsive system can be seen as long-term memory and therefore has functionally unlimited capacity. The impulsive system forms associative links between individual elements that may be of varying strengths. When an element is activated, other elements linked to it are activated in accordance with the strength of the links to the original element, a form of spreading activation. Recent or frequent activation of an element also increases both the accessibility of that element and the likelihood of its continued processing if further activation occurs. Links between elements are based on the principles of contiguity and similarity, so that stimuli that occur a short time or distance from one another are more likely to become linked or to have existing links strengthened than is the case with temporally or spatially nonproximal stimuli. It is important to note that reflective processes activate corresponding patterns of impulsive elements, meaning that even elements that are never perceived together in the world may become associatively linked if they are often reflectively processed together. The activation of impulsive elements may be accompanied by an experiential state of awareness, with people experiencing a specific feeling without being able to say where it comes from. Such feelings are themselves elements in the impulsive system and may therefore be connected to other elements with varying link strengths.

This process of association is slow and enduring but also rigid. Although propositional processing in the reflective system may create associative links between concepts in the impulsive system, it is not necessarily the case that propositional

processing can become completely automatic and effortless with enough practice (Deutsch et al., 2009). Instead, specific patterns of thought may become easier but lose or change their meaning under circumstances where reflective processing is hindered. As an example, African Americans who feel that whites perceive them negatively may often reflectively think "Blacks are not bad" as a response; in the impulsive system, however, the concepts of *black* and *bad* are being activated at the same time and thereby linked (Livingston, 2002). This simultaneity may counterintuitively lead African Americans to have additional negative feelings and associations with their in-group when they have no reflective resources with which to negate the associative relation. However, the rigidity of associative processing is somewhat alleviated by the effects of motivational orientation. When people are in an approach orientation, they are prepared to reduce the distance between themselves and some aspect of their environments, whereas an avoidance orientation is preparedness to increase this distance. These fundamental orientations can facilitate associative activation when this activation is compatible (e.g., approach orientation may facilitate the activation of positively valent elements). Therefore, the sight of a muffin might elicit the positive associations of *tasty*, *sweet*, and *good* more strongly when one is in an approach orientation, as when selecting what salad one would prefer for lunch, than when one is not. If one is currently in an avoidance orientation, having just read a brochure on calorie content and having selected the salad that would be least fattening, these associations might be weaker even before the activation of the reflective system. However, it is clear that the associative system alone, although it requires no resources and works very quickly, is generally far less flexible and adaptive than the reflective system.

## *Interaction of Systems*

Now that we have described the basic operating principles of the RIM's two systems, it is necessary to show how they function together. The very structures of the two systems contain a first important point relevant to their interaction: the impulsive system is always effortlessly active, whereas the reflective system may also be inactive. The implication is that the reflective system, when it does operate, does so in parallel with the impulsive system, not in place of it. That is, reflective processing always occurs with parallel impulsive processing. It is also clear that the concepts that are transformed in working-memory space in the reflective system do not come from nowhere but from the long-term store of the impulsive system.

As the systems cannot interact when the reflective system is disengaged, it is adequate to examine how they interact from the beginning to the end of a reflective operation. When a reflective operation begins, perceptual input will already have activated several associative elements. For example, when thinking about what to have for lunch, a person may already have seen what is on offer in the cafeteria, a selection that will activate whatever associations that person has with the given meal options, but other perceptual data in the attentional focus (the presentation of the

food, the attractiveness of the serving staff, and any number of other concepts) will also activate associations of their own. Some of these associations will achieve sufficient activation to attain awareness, so a particularly delicious-looking muffin might prompt an automatic expectation of a good taste. The reflective system then categorizes and relates the activated concepts, the result being that the muffin is recognized as tasting good, and activates additional relevant content in the impulsive system—such as health. This concept, in turn, changes the activation pattern in the impulsive system, so the associated concept of *salad* might become activated as well. This activation pattern is again categorized, and the process repeats until a decision or inference is reached. Such end results in the reflective system are driven by the principle of consistency of the propositions generated. For example *health is good*, *taste is good*, *health is more important than taste* might lead to the decision to select the healthy salad rather than the unhealthy one but also the tasty muffin, for this choice would be consistent with a greater number of propositions. However, the content of the propositions generated is necessarily limited by the activation pattern of the impulsive system—although tennis is also healthy, its activation potential in the environment of the cafeteria during lunch hour is very low, so the reflective system will not include it in processing without any prior link or further relevant perceptual stimuli.

Synergy between the systems occurs when the impulsive system's associations are valid and relevant to a consistent reflective solution. When the impulsive activation pattern is in synergy with reflective processing, concepts relevant to the focus of reflective processing become comparatively accessible, and cognitive effort is therefore reduced. The reflective system is not forced to perform extra categorizations and activations of concepts to achieve consistency, so subjective effort is lessened. This reduction may be accompanied by a feeling of flow (Winkielman, Huber, Kavanagh, & Schwarz, 2012), that is, ease of processing, which is then linked to positive affect. Therefore, when both systems are in accord, it feels easy and good to think and make decisions. As an example, if the only tasty option in the cafeteria were the healthy salad, people who ate there and cared deeply about their health (i.e., had a high accessibility of the concept health) would find it natural to choose the salad and, moreover, would feel good about how easy the choice was. A different picture emerges when the systems are at odds with each other, as when impulsive activation patterns present associations that are opposed to a consistent reflective conclusion and produce a feeling of conflict. It requires additional cognitive effort to activate new impulsive patterns and to form propositions that lead to a consistent end state. Once the muffin is added to the lunch options, the decision-maker must actively work against the temptation of the tasty dessert in order to generate the propositions about healthy eating that justify selecting the salad. This dependence of effortful processing on automatic activation has an interesting consequence: A fluently (synergistically) processed inference should have a higher truth value than a disfluently (antagonistically) processed inference does, unless the reflective system specifically corrects for the consequences of fluency (Allport & Lepkin, 1945; Begg, Anas, & Farinacci, 1992; Schwarz, Sanna, Skurnik, & Yoon, 2007). The

fluency of processing should affect both how sure a person is of a syllogistic inference and how securely she or he stands behind a given decision.

## Common Pathway to Behavior

Having described how the reflective and impulsive systems are structured and how they interact during the process of thought, we now turn to two issues that many other dual-process models have not yet addressed: how these mental processes are linked to behavior and especially how they interact when causing behavior. The RIM seeks to provide an answer to this central question through the component of behavioral schemata. In a general sense, behavioral schemata are clusters of elements in the impulsive system. They consist of specific motor representations of behavior, the perceptual input of typical context factors for the behavior, and the consequences of those factors. As elements of the impulsive system, specific behavioral schemata (e.g., how to grip a fork) will likely be associated with other behavioral schemata that are relevant to the context, the specific motor activation, or the consequences of the behavior. This result leads to the conclusion that several specific, concrete behavioral schemata might form clusters that, in turn, can be seen as behavioral schemata of greater abstractness. Gripping a fork and gripping a spoon, for instance, are similar in terms of expected consequences (tasting food) and context (involving food, eating utensils, etc.). The more abstract behavioral schema in this case might be termed *holding cutlery*, which might, in turn, connect to other schemata to form a cluster of *eating politely*, and so on. This conceptual hierarchy bears some relationship to other conceptualizations of schemata but is more specific than its typical use (for a discussion see Fiske & Linville, 1980; Lodge, McGraw, Conover, Feldman, & Miller, 1991).

If behavioral schemata are situated in the impulsive system, they are also subject to spreading activation. In a way similar to that of nonbehavioral elements in the impulsive system, behavioral schemata can be activated automatically if enough other elements with sufficiently strong links to the schema are activated. In the case of behavioral schemata, activation includes performance of the motor program embedded within the schema. The reflective system, on the other hand, activates behavioral schemata by propositionally connecting the self to the required behavioral schema by means of a behavioral decision. If the decision refers to a behavior that is to take place in future, the reflective system additionally links the behavior's expected relevant context with the cluster, creating an intention (Gollwitzer, 1999). Although the systems encompass different precursors to behavior, they use the same final gateway to enact behavior, namely, the activation of behavioral schemata, elements of the impulsive system. The implication is that the same principles of system interaction that apply to purely mental operations in the RIM also apply to behavioral processes—fluent, synergistic processing of behavioral decisions leads to behavior that is performed more smoothly and easily than disfluent, antagonistic processing. In extreme cases of the latter, the impulsive activation of behavioral

schemata might be so strong that a reflective behavioral decision will not lead directly to behavior, as when one knows that eating the dessert is wrong and has decided not to, but the eyes and hands still seem to have their own volition to fixate on the forbidden. In this case the operating conditions of the systems will decide which behavior is performed. In states inimical to reflective processing, such as high arousal or reduced working memory capacity, the impulse to act will likely win out, whereas an unhindered reflective system might redirect the activation pattern in the impulsive system by, for example, diverting attention or actively linking negative consequences or attributes to the impulsive behavior. In this sense, there is little difference between mentally grappling with an unwanted impulsive idea, such as unwanted racist thoughts, and with an unwanted physical impulse, such as reaching for that tempting muffin.

## *Knowledge and Action: Bidirectional Connections*

Having established that behavioral schemata are situated in the impulsive system, we note that an important additional implication of the RIM becomes clear. In the impulsive system, each element adheres to the mechanism of spreading activation, including behavioral schemata. Because behavioral schemata are linked to consequences and contextual stimuli, activating the behavior also activates related concepts. Therefore, behavior can directly influence cognition, both by creating and modifying associative links and by influencing what concepts are likely to become active in the reflective system. For example, the act of reaching for the muffin, if repeated often, may eventually become sufficiently associated with the sight of the muffin that the action is triggered by the sight. Similarly, this association makes the muffin more likely to come to mind when one is reaching for the salad, that is, when performing the very motor program associated with the muffin. This bidirectionality of influence—behavior to cognition as well as cognition to behavior—also holds for motivational orientation in that behaviors associated with approach will activate an approach orientation, whereas avoidance behavior has the opposite effect.

This link between behavior and cognition is not a new idea in principle. A similar effect is assumed in several influential theories, such as cognitive dissonance theory (Festinger, 1957) and self-perception theory (Bem, 1967). They also predict a change in attitudes as a specific form of cognition that is based on behavior, albeit via different mechanisms. The important difference lies in precisely these mechanisms. According to cognitive dissonance theory, a behavior that runs counter to an existing belief about the self causes an aversive motivational state that may be alleviated by changing existing beliefs or adding new ones, whereas the assumption in self-perception theory is that people infer their attitudes toward particular objects from their own behavior toward those objects. Both of these proposed mechanisms require propositional processing because they depend on a categorization of the behavior in question. The RIM, on the other hand, describes a direct link between behavior and cognition via the associative links between behavioral schemata and

contiguously activated concepts in the impulsive system. It is not that cognitive dissonance and self-perception theory lack validity but rather that their path from behavior to cognitive change is not the only possible one.

When this perspective on the connection between behavior and implicit associative links is applied to the field of implicit attitudes, several interesting implications arise. From the perspective of the RIM, implicit attitudes measured by the IAT, affective priming, and other procedures based on reaction time reflect the strength of associative links between a target and a valence by means of a behavior associated with that valence.[1] In view of the attributes of the impulsive system, it follows that negated targets or valences should, under certain circumstances, have the same effect on reaction times as nonnegated ones do. Evidence supporting this logic comes from the Bona Fide Pipeline task (Fazio, Jackson, Dunton, & Williams, 1995), which was used in a study by Deutsch, Gawronski, and Strack (2006), who showed that positive prime words facilitated categorization of positive targets even when the prime was negated, whereas both negated and nonnegated negative prime words facilitated categorization of negative target. Although this effect is not universal (Deutsch et al., 2009), it provides evidence that impulsive associations are what implicit attitude procedures measure. Procedures used to change implicit attitudes, such as automatic stereotype reduction training (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000), can also benefit from the implications of the RIM's structure. In the original automatic stereotype reduction training, participants were induced to respond to stereotype-congruent pairings with a NO key and to stereotype-incongruent pairings with a YES key. Although this method was effective in reducing automatic stereotype activation, a study using only stereotype-incongruent pairings has shown the procedure to work (Gawronski, Deutsch, Mbirkou, Seibt, & Strack, 2008). The same studies also showed that negation of stereotype-congruent pairings alone actually increased implicit stereotyping, a prediction derived from the contrast between the propositional nature of negation and the associative nature of contiguous stimuli pairs.

The bidirectional link between behavior and cognition in the RIM also makes that model a valuable framework for studies that deal with embodied effects on cognition. The basic idea of embodiment research is that cognition is always founded in mechanisms of sensory processing and motor control (e.g., Wilson, 2002). The role of perceptual information and motor-processing in the impulsive system and the influence thereby exerted on the reflective system mirror this basic assumption. Precursor research to the current wave of embodiment movement, such as the pen study by Strack, Martin, and Stepper (1988) and the headphone study by Förster and Strack (1996), have already been discussed from the perspective of the RIM (Strack & Deutsch, 2004), but newer investigations into embodied processes might also be integrated into this model. For example, Zhong and Lijenquist (2006) show that washing hands reduced the effects of guilt on compensatory prosocial behavior. The authors explain the effect in terms of symbolic self-completion (Wicklund & Gollwitzer, 1981),

[1] The IAT is also capable of measuring associations other than valence, but because the argument is analogous to valence associations, it is omitted here.

but an alternative explanation offered by the RIM would be that the metaphorical association of purity with cleanliness might conflict with a negative activation of the self in the context of guilt and thereby alleviate the guilt's effects on behavior. Thus, an impetus is given for further inquiries into the precise mechanism behind embodiment effects.

## *Determinants of Systematic Behavioral Control*

Having established the parameters and attributes of the two systems of the RIM and their parallel effects on behavior, we turn to explaining possible determinants of either system's dominance over the other in having these effects. If it is possible to identify variables that can affect which system determines behavior, one can vastly improve the predictive power of the RIM for behavior. Because the impulsive system is always active and processing, this question can be rephrased: Under what circumstances will the reflective system assert behavioral control when in conflict with the impulsive system?

### Motivation and Opportunity

Fazio (1990) describes two modes of thinking—a spontaneous processing mode based on attitude accessibility, and a deliberative processing mode based on attitude behavior. These modes of thinking are remarkably similar in structure to the RIM's proposed systems, certain differences in mechanisms notwithstanding. Therefore, the MODE model (Fazio, 1990), which predicts when the deliberative mode will be engaged in processing the possible consequences of behavior, may be applicable to the RIM as well. In this conceptualization, engagement in deliberative processing depends on motivation and opportunity. Motivation in the MODE model is generated by the fear of invalidity (Kruglanski & Freund, 1983), a function of the perceived costliness of a judgmental mistake to the self, whereas opportunity is a function of the available time and resources for processing. Applying this framework to the RIM, one finds that the defined properties of the reflective system are in accord with these predictions. Reflective processing is accompanied by a feeling of subjective effort and so requires motivation, whereas both the reliance on the resources of working memory and the relatively slow speed of the reflective system make it clear that the reflective system can influence behavior only if the opportunity is given.

Evidence for this dependence on opportunity exists in many domains. Cognitive load, a manipulation often used to impair deliberative processing, has been applied in various different studies whose results can be explained with the RIM. Self-control (e.g., Lattimore & Maxwell, 2004; Wegner, Erber, & Zanakos, 1993), processing of negated stimuli (Deutsch et al. 2009), social judgments and attributions (Gilbert et al., 1988; Krull & Erickson, 1995; Trope & Alfieri, 1997), moral judgments (Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008), and general reasoning (De Neys, 2006) have all proven to be impaired by cognitive load in ways

that are consistent with the RIM's predictions. The idea that working memory resources may also play a role in reflective processing has been tested in several studies, both by comparing individuals with dispositionally high or low working memory capacities (Hofmann, Gschwendner, Friese, Wiers, & Schmitt, 2008; Thush et al., 2008) and by specifically taxing resources of working memory (Deutsch et al., 2009). The conceptualization of working memory as "a domain-free limitation in ability to control attention" (Engle, 2002, p. 19) points to the conclusion that the effects of attentional cognitive load manipulations on reflective processing may be mediated by working memory capacity.

### Self-Regulatory Resources

Vohs (2006) argues that the RIM's reflective system is, in fact, powered by self-regulatory resources (Baumeister, Bratslavsky, Muraven, & Tice, 1998). This account is based on the similarity between the idea of such resources and the cognitive resources discussed in relation to the RIM. Vohs & Faber, (2007) argue that impulsive spending, a phenomenon expected to occur under circumstances that inhibit reflective processing, does indeed occur more often when participants are depleted of self-regulatory resources. A structurally similar argument is made with respect to overeating among dieters (Vohs & Heatherton 2000). Further evidence for this integration of theories comes from research showing that effortful self-regulation has a detrimental effect on subsequent reasoning capabilities (Schmeichel, Vohs, & Baumeister, 2003). The bidirectionality of this effect, that is, a negative effect that prior reflective decision-making exerts on subsequent self-regulation, is shown in a laboratory paradigm encompassing many simple decisions followed by a self-regulatory exercise. The effect's bidirectionality also surfaces in a field study in which shoppers who reported having made effortful decisions previously solved fewer math problems than those who had engaged in fewer decisions during their shopping trip (Vohs et al., 2008). Although this evidence hints at a connection between self-regulatory and reflective cognitive resources, these studies do not show a direct link between the two. Other research shows that dietary standards and explicit target attitudes predict behavior only when self-regulatory resources are available; when it is not, implicit attitudes are better predictors (Friese, Hofmann, & Wänke, 2008; Hofmann, Rauch, & Gawronski, 2007). The fact that impulsive and reflective predictors diverge depending on the availability of self-regulatory resources underlines the conceptualization of self-regulation as a conflict between impulsive and reflective behavioral activation. Together with the evidence presented by Vohs (2006), these findings permit the conclusion that research on resource-based self-regulation can be integrated into the RIM. It remains to be seen whether self-regulatory resources are equivalent to working memory resources or whether they constitute their own construct.

## Implications of the RIM

### *Thinking Is Tough!*

Coming up with a complex plan of action is not a simple endeavor. The Trojan Horse required the cunning Odysseus to think hard for a long time, and this story is one of the main reasons he endures as a hero figure. In general terms, it is not easy to engage in reflective processing—beyond the subjective feeling of difficulty, there may be physical limits to the human ability to think (Gailliot et al., 2007). Although thinking may sometimes be facilitated when the reflective and impulsive systems are in accord, people must often use reflective processing against the pull of impulsive associations. Whether this struggle is due to temptation or to particularly complex challenges in the environment, the difficulty in staying the reflective course is clear.

However, cultivating moral or thoughtful habits may become easier with time. Specific propositional operations can become associated with the feeling of temptation if they are activated often enough, and even the experiential component of reflective operation (the feeling of effortful cognition) is itself represented in the impulsive system and may thus become associated with it. For careful planning habits, positive affect associated with successful plans may lead to the process of planning itself acquiring a positive valence, with these habits of thought perhaps eventually becoming inculcated through successful implementation. However, this effect is not sufficient to become truly automatic. Although reflective processing may become facilitated by such mechanisms, propositional reasoning itself cannot become automated. If resources are lacking, not even these habits of reflection will make for better control of impulses or careful planning. No matter how accessible the relevant propositional transformations may be in the impulsive system, reflective resources are required if a person is actually to bring those transformations to bear upon activated concepts.

The habit of critical metacognition is a particularly interesting case. Metacognition refers to thoughts about one's own thoughts, and critical metacognition is therefore those thoughts that evaluate the thinking process. In cognitive-behavioral therapy, thoughts are actively evaluated by the patient and classified as rational or irrational (Baer, 2003). The goal of such interventions is often to change dysfunctional behavior or thinking patterns, such as "catastrophizing" (Beck, 1976). Pursuing this kind of metacognitive thought alteration or suppression may be especially difficult because of the vast reflective resources required. A strong association between the metacognitive monitoring process and the undesired thoughts would eventually activate the latter rather than suppress them, requiring additional reflective resources to eliminate them. Evidence from studies on emotional disorders shows that metacognitive thought suppression does indeed increase the frequency of unwanted thoughts (Purdon, 1999). Although long-term use of metacognitive strategies may eventually divest undesired thoughts of their potency, it seems clear that the way there is a long and cognitively taxing one.

## *Sometimes No Means Yes If I Can't Process It*

Given the research showing that negation is a reflective process that requires reflective resources (Deutsch et al., 2006, 2009; Grant, Malaviya, & Sternthal, 2004), one must wonder at the efficacy of negated persuasive messages. Evidence showing that older adults may be more likely to misremember explicitly negated information has been interpreted as a warning against negatively worded statements in healthcare materials (Wilson & Park, 2008). Wakefield et al.'s (2008) finding that antismoking campaigns funded by tobacco companies have few, if any, negative effects on teenagers' intentions to smoke may be partially due to the negated messages employed (e.g., "Think, Don't Smoke"; see Farrelly, Niederdeppe, & Yarsevich, 2003). These effects can be explained by a lack of processing resources in recalling the content of the messages. If advertisements or healthcare instructions convey associations that are negated in the message, then recall of the message may be confined to the association between the elements under circumstances of low reflective resources. This limitation can lead negated statements to be remembered as affirmed. Positively formulated messages (i.e., those whose association mirrors their propositional connection) are more effective, as with antismoking ads that depict smoking in combination with serious consequences for health (Biener, 2002).

This logic might also be generalized to ethical norms in everyday life. If presented as negated statements (such as *don't drink and drive*), strong ethical norms may have a behavioral effect opposite to what is expected. If people are continually reminded of what they should not do, the unwanted action will be continually activated in the impulsive system and thereby affect subsequent reflective and impulsive operations. If a sermon on the forbidden practice of adultery activates that concept and its influence on evaluations, the listener could see some other members of the congregation as attractive for possibly committing adultery with them. Of course, this possibility would be negated, but only if the listener had sufficient reflective resources available to perform the negation. It would be particularly problematic in cases where the forbidden behavior is hedonically attractive to start with; the activation of the concept would then also activate the expectation of the pleasant feeling associated with it and elicit an approach orientation. A better approach to ethics according to this logic might instead be to praise exemplars of morality or to prescribe morally positive acts as opposed to forbidding negative ones. Instead of saying no to vice, one should say yes to virtue.

## *Improving Implicit Self-Esteem*

Implicit self-esteem is a much researched construct of recent years (e.g., Koole, Dijksterhuis, & van Knippenberg, 2001; Yamaguchi et al., 2007). In a general sense, implicit self-esteem is defined as an automatic evaluation of the self that occurs nonconsciously and affects spontaneous reactions to self-relevant stimuli (Bosson, et al. 2000). Implicit self-esteem has been shown to specifically predict diverse

outcomes in a range of situations, including apparent anxiety in participants as they complete a self-relevant interview (Spalding & Hardin, 1999), levels of implicit gender bias in combination with implicit gender identity (Aidman & Carroll, 2003), and depressive symptomatology 6 months after measurement (Franck, De Readt, & De Houwer, 2007). Explicit self-esteem (i.e., explicit evaluations of self-worth), on the other hand, has specific predictive power for subjective well-being (Schimmack & Diener, 2003). It is the combination of the two constructs, however, that has excited most interest in recent years. In particular, the question of discrepant self-esteem has been examined (Zeigler-Hill, 2006). That is, what effects do high explicit but low implicit, or low explicit but high implicit, self-esteem have on behavior? One direct prediction based on psychodynamic theory concerns narcissism. The *mask model* of narcissism assumes that narcissistic individuals are characterized by deep self-doubt (corresponding to low implicit self-esteem), which they compensate for by projecting grandiose self-views (corresponding to extremely high explicit self-esteem) (Bosson et al., 2008). Empirical evidence supporting this model has remained mixed, however, perhaps partly because of the relative unreliability of measures of implicit self-esteem (Bosson et al., 2000).

By rather precisely spelling out the concepts involved, the RIM can contribute to the discussion of these and other phenomena of the implicit self. Implicit self-esteem might, for instance, be conceptualized as the total valence of the associative pattern linked with activation of the self in the impulsive system. This specification would imply recommendations for effective measures of implicit self-esteem—namely, measures that do not require any explicit judgment but rather depend solely on valence and behavior interference (e.g., the IAT, measures based on the logic of affective priming). In addition, the same logic that is applied to automatic stereotype reduction training (Gawronski et al., 2008) may be applied to the implicit self-esteem construct. Doing so yields a method through which to increase implicit self-esteem by consistently affirming positive pairings of valence and self (e.g., Dijksterhuis, 2004), suggesting a possible avenue for therapy of narcissism and other negative effects of low implicit self-esteem.

## Conclusion

The RIM offers a multitude of predictions that can help improve the understanding of the link between knowledge and action, whether it be explaining the reasoning processes behind complex plans such as the Trojan Horse, the seemingly self-destructive flight of Icarus, or even the conflict between rationality and impulse as epitomized in Odysseus' suffering of the Siren's song. Although effortful, reflective processing may occur in fluent synergy with impulsive processing, there are often conflicts between the two systems. Their resolution is a question of available reflective resources and motivation to use them. But whether the systems work in concert or struggle against one another, the pathway to behavior is ultimately the same—behavioral schemata are activated depending on the results of both

systems' processing. However, it is not just knowledge that may determine action in the context of the RIM. The bidirectional associative links between behavioral schemata and other elements of the impulsive system mean that action, in turn, influences knowledge. This bidirectionality in combination with the functioning of the two systems can tell much about what it means to think and act, when No might appear to mean Yes, and how to pinpoint and modify the elusive implicit self.

## References

Aidman, E. V., & Carroll, S. M. (2003). Implicit individual differences: Relationships between implicit self-esteem, gender identity, and gender attitudes. *European Journal of Personality, 17,* 19–37. doi:10.1002/per.465

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In P. D. J. Kuhl & D. J. Beckmann (Eds.), *Action control* (pp. 11–39). Berlin: Springer. doi:10.1007/978-3-642-69746-3_2

Allport, F. H., & Lepkin, M. (1945). Wartime rumors of waste and special privilege: Why some people believe them. *The Journal of Abnormal and Social Psychology, 40,* 3–36. doi:10.1037/h0058110

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40,* 471–499. doi:10.1348/014466601164939

Baddeley, A. (1986). Working memory, reading and dyslexia. In E. Hjelmquist & L. G. Nilsson (Eds.), *Communication and handicap* (pp. 141–152). Advances in Psychology: Vol. 34. North-Holland: Elsevier Science. doi:10.1016/S0166-4115(08)61202-9

Baer, R. A. (2003). Mindfulness training as a clinical intervention: A conceptual and empirical review. *Clinical Psychology: Science and Practice, 10,* 125–143. doi:10.1093/clipsy.bpg015

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Experimental Psychology, 48,* 145–160. doi:10.1026//0949-3946.48.2.145

Bargh, J. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *The automaticity of everyday life* (pp. 1–61). Advances in Social Cognition: Vol. 10. Mahwah: Lawrence Erlbaum.

Baron, R. S. (2000). Arousal, capacity, and intense indoctrination. *Personality and Social Psychology Review, 4,* 238–254. doi:10.1207/S15327957PSPR0403_3

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74,* 1252–1265. doi:10.1037/0022-3514.74.5.1252

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. Madison: International Universities Press.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121,* 446–458. doi:10.1037/0096-3445.121.4.446

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review, 74,* 183–200. doi:10.1037/h0024835

Biener, L. (2002). Anti-tobacco advertisements by Massachusetts and Philip Morris: What teenagers think. *Tobacco Control,* 11(S2), ii43–ii46. doi:10.1136/tc.11.suppl_2.ii43

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Development of cognition, affect, and social relations* (pp. 39–101). The Minnesota Symposia on Child Psychology: Vol. 13. Hillsdale: Erlbaum.

Bosson, J. K., Lakey, C. E., Campbell, W. K., Zeigler-Hill, V., Jordan, C. H., & Kernis, M. H. (2008). Untangling the links between narcissism and self-esteem: A theoretical and empirical

review. *Social and Personality Psychology Compass, 2,* 1415–1439. doi:10.1111/j.1751-9004.2008.00089.x

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631–643. doi:10.1037/0022-3514.79.4.631

Bushman, B. J., & Cooper, H. M. (1990). Effects of alcohol on human aggression: An integrative research review. *Psychological Bulletin, 107,* 341–354. doi:10.1037/0033-2909.107.3.341

Carver, C. S. (2005). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review, 9,* 312–333. doi:10.1207/s15327957pspr0904_2

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (pp. 3–39). The Ontario Symposium: Vol. 5. Hillsdale: Erlbaum.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Conner, M., & Armitage, C. J. (1998). Extending the theory of planned behavior: A review and avenues for further research. *Journal of Applied Social Psychology, 28,* 1429–1464. doi:10.1111/j.1559-1816.1998.tb01685.x

Crisp, R. (2000). *Aristotle: Nicomachean ethics*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511802058

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17,* 428–433. doi:10.1111/j.1467-9280.2006.01723.x

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91,* 385–405. doi:10.1037/0022-3514.91.3.385

Deutsch, R., Kordts-Freudinger, R., Gawronski, B., & Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology, 56,* 434–446. doi:10.1027/1618-3169.56.6.434

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. doi:10.1037/0022-3514.56.1.5

DeWall, C. N., Baumeister, R. F., & Masicampo, E. J. (2008). Evidence that logical reasoning depends on conscious processing. *Consciousness and Cognition, 17,* 628–645. doi:10.1016/j.concog.2007.12.004

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86,* 345–355. doi:10.1037/0022-3514.86.2.345

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth: Harcourt Brace.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23. doi:10.1111/1467-8721.00160

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62,* 328–339. doi:10.1037/0022-3514.62.2.328

Farrelly, M. C., Niederdeppe, J., & Yarsevich, J. (2003). Youth tobacco prevention mass media campaigns: Past, present, and future directions. *Tobacco Control, 12*(S1), i35–i47. doi:10.1136/tc.12.suppl_1.i35

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York: Academic Press.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion, 15,* 115–141. doi:10.1080/02699930125908

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027. doi:10.1037/0022-3514.69.6.1013

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54,* 297–327. doi:10.1146/annurev.psych.54.101601.145225

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238. doi:10.1037/0022-3514.50.2.229

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology, 27,* 307–316. doi:10.1207/s15324834basp2704_3

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research.* Reading: Addison-Wesley.

Fiske, S. T., & Linville, P. W. (1980). What does the schema concept buy us? *Personality and Social Psychology Bulletin, 6,* 543–557. doi:10.1177/014616728064006

Förster, J., & Strack, F. (1996). Influence of overt head movements on memory for valenced words: A case of conceptual-motor compatibility. *Journal of Personality and Social Psychology, 71,* 421–430. doi:10.1037/0022-3514.71.3.421

Franck, E., De Raedt, R., & De Houwer, J. (2007). Implicit but not explicit self-esteem predicts future depressive symptomatology. *Behaviour Research and Therapy, 45,* 2448–2455. doi:10.1016/j.brat.2007.01.008

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology, 47,* 397–419. doi:10.1348/014466607X241540

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology, 92,* 325–336. doi:10.1037/0022-3514.92.2.325

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When "Just Say No" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44,* 370–377. doi:10.1016/j.jesp.2006.12.004

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40,* 535–542. doi:10.1016/j.jesp.2003.10.005

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119. doi:10.1037/0003-066X.46.2.107

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54,* 733–740. doi:10.1037/0022-3514.54.5.733

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54,* 493–503. doi:10.1037/0003-066X.54.7.493

Grant, S. J., Malaviya, P., & Sternthal, B. (2004). The influence of negation on product evaluations. *Journal of Consumer Research, 31,* 583–591. doi:10.1086/425093

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107,* 1144–1154. doi:10.1016/j.cognition.2007.11.004

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27. doi:10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480. doi:10.1037/0022-3514.74.6.1464

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science, 4,* 162–176. doi:10.1111/j.1745-6924.2009.01116.x

Hofmann, W., Gschwendner, T., Friese, M., Wiers, R. W., & Schmitt, M. (2008). Working memory capacity and self-regulatory behavior: Toward an individual differences perspective on behav-

ior determination by automatic versus controlled processes. *Journal of Personality and Social Psychology, 95,* 962–977. doi:10.1037/a0012705

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology, 43,* 497–504. doi:10.1016/j.jesp.2006.05.004

James, W. (1890). *Principles of psychology*. New York: Holt.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78,* 871–888. doi:10.1037/0022-3514.78.5.871

Kim, D.-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66,* 83–96. doi:10.2307/3090143

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80,* 669–685. doi:10.1037//0022-3514.80.4.669

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19,* 448–468. doi:10.1016/0022-1031(83)90022-7

Krull, D. S., & Erickson, D. J. (1995). Judging situations: On the effortful process of taking dispositional information into account. *Social Cognition, 13,* 417–438. doi:10.1521/soco.1995.13.4.417

Lambert, A. J., Payne, K. B., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology, 84,* 277–295. doi:10.1037/0022-3514.84.2.277

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36,* 635–642. doi:10.1037/0022-3514.36.6.635

Lattimore, P., & Maxwell, L. (2004). Cognitive load, stress, and disinhibited eating. *Eating Behaviors, 5,* 315–324. doi:10.1016/j.eatbeh.2004.04.009

Livingston, R. W. (2002). The role of perceived negativity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology, 38,* 405–413. doi:10.1016/S0022-1031(02)00002-1

Lodge, M., McGraw, K. M., Conover, P. J., Feldman, S., & Miller, A. H. (1991). Where is the schema? Critiques. *The American Political Science Review, 85,* 1357–1380. doi:10.2307/1963950

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Experimental Psychology, 48,* 107–122. doi:10.1026//0949-3946.48.2.107

Mischel, W., Ebbesen, E. B., & Raskoff, Z. A. (1972). Cognitive and attentional mechanisms in the delay of gratification. *Journal of Personality and Social Psychology, 21,* 204–218. doi:10.1037/h0032198

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126,* 247–259. http://dx.doi:10.1037//0033-2909.126.2.247

Paulhus, D. L., & Lim, D. T. K. (1994). Arousal and evaluative extremity in social judgments: A dynamic complexity model. *European Journal of Social Psychology, 24,* 89–99. doi:10.1002/ejsp.2420240107

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer.

Purdon, C. (1999). Thought suppression and psychopathology. *Behaviour Research and Therapy, 37,* 1029–1054. doi:10.1016/S0005-7967(98)00200-9

Schachter, S. (1971). Some extraordinary facts about obese humans and rats. *American Psychologist, 26,* 129–144. doi:10.1037/h0030817

Schimmack, U., & Diener, E. (2003). Predictive validity of explicit and implicit self-esteem for subjective well-being. *Journal of Research in Personality, 37,* 100–106. doi:10.1016/S0092-6566(02)00532-9

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology, 85,* 33–46. doi:10.1037/0022-3514.85.1.33

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 39, pp. 127–161). New York: Academic Press. doi:10.1016/S0065-2601(06)39003-X

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22. doi:10.1037/0033-2909.119.1.3

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108–131. doi:10.1207/S15327957PSPR0402_01

Spalding, L. R., & Hardin, C. D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10,* 535–539. doi:10.1111/1467-9280.00202

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247. doi:10.1207/s15327957pspr0803_1

Strack, F., & Deutsch, R. (2007). The role of impulse in social behavior. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (revised and expanded 2nd ed., pp. 408–431). New York: Guilford Press.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology, 54,* 768–777. doi:10.1037/0022-3514.54.5.768

Strack, F., & Neumann, R. (1996). "The spirit is willing, but the flesh is weak": Beyond mind-body interaction in human decision-making. *Organizational Behavior and Human Decision Processes, 65,* 300–304. doi:10.1006/obhd.1996.0031

Sutton, S., Marsh, A., & Matheson, J. (1987). Explaining smokers' decisions to stop: Test of an expectancy-value approach. *Social Behaviour, 2,* 35–49.

Thush, C., Wiers, R. W., Ames, S. L., Grenard, J. L., Sussman, S., & Stacy, A. W. (2008). Interactions between implicit and explicit cognition and working memory capacity in the prediction of alcohol use in at-risk adolescents. *Drug and Alcohol Dependence, 94,* 116–124. doi:10.1016/j.drugalcdep.2007.10.019

Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology, 73,* 662–674. doi:10.1037/0022-3514.73.4.662

Vohs, K. D. (2006). Self-regulatory resources power the reflective system: Evidence from five domains. *Journal of Consumer Psychology, 16,* 217–223. doi:10.1207/s15327663jcp1603_3

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making choices impairs subsequent self-control: A limited-resource account of decision making, self-regulation, and active initiative. *Journal of Personality and Social Psychology, 94,* 883–898. doi:10.1037/0022-3514.94.5.883

Vohs, K. D., & Faber, R. J. (2007). Spent resources: Self-regulatory resource availability affects impulse buying. *Journal of Consumer Research, 33,* 537–547. doi:10.1086/510228

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science, 11,* 249–254. doi:10.1111/1467-9280.00250

Vroom, H, V. (1964). *Work and motivation*. New York: Wiley.

Wakefield, M. A., Durkin, S., Spittal, M. J., Siahpush, M., Scollo, M., Simpson, J. A., Chapman, S., White, V., & Hill, D. (2008). Impact of tobacco control policies and mass media campaigns on monthly adult smoking prevalence. *American Journal of Public Health, 98,* 1443–1450. doi:10.2105/AJPH.2007.128991

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology, 11,* 92–107. doi:10.1080/17470215908416296

Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology, 65,* 1093–1104. doi:10.1037/0022-3514.65.6.1093

Wicklund, R. A., & Gollwitzer, P. M. (1981). Symbolic self-completion, attempted influence, and self-deprecation. *Basic and Applied Social Psychology, 2,* 89–114. doi:10.1207/s15324834basp0202_2

Wigfield, A., & Eccles, J. S. (2000). Expectancy—Value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68–81. doi:10.1006/ceps.1999.1015

Wilson, E. A. H., & Park, D. C. (2008). A case for clarity in the writing of health statements. *Patient Education and Counseling, 72,* 330–335. doi:10.1016/j.pec.2008.02.008

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636. doi:10.3758/BF03196322

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103–128. doi:10.1016/0010-0277(83)90004-5

Winkielman, P., Huber, D. E., Kavanagh, L., & Schwarz, N. (2012). Fluency of consistency: When thoughts fit nicely and flow smoothly. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 89–111). New York: Guilford Press.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827. doi:10.1037/0022-3514.81.5.815

Yamaguchi, S., Greenwald, A. G., Banaji, M. R., Murakami, F., Chen, D., Shiomura, K., Kobayashi, C., Cai, H., & Krendl, K. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science, 18,* 498–500. doi:10.1111/j.1467-9280.2007.01928.x

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459–482. doi:10.1002/cne.920180503

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175. doi:10.1037/0003-066X.35.2.151

Zeigler-Hill, V. (2006). Discrepancies between implicit and explicit self-esteem: Implications for narcissism and self-esteem instability. *Journal of Personality, 74,* 119–144. doi:10.1111/j.1467-6494.2005.00371.x

Zhong, C.-B., & Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science, 313,* 1451–1452. doi:10.1126/science.1130726